

Nat Lang Linguist Theory (2009) 27: 197–224
DOI 10.1007/s11049-008-9061-1

ORIGINAL PAPER

Diverse repairs for multiple labial consonants

Kie Zuraw · Yu-An Lu

Received: 19 February 2008 / Accepted: 30 November 2008 / Published online: 17 February 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract The relationship between constraints on surface forms and operations that alter representations is of central interest in phonological theory. This squib presents a case of diverse “repairs” in response to a marked structure—labial ... labial sequences—created by *um*-infixation in stems beginning with (or, in some cases, merely containing) labial consonants in Austronesian languages. We review several strategies, which for the most part do not cluster according to subfamilies: tolerance, gaps, loss of stem consonant, loss of infix nasality, stem dissimilation, infix dissimilation, prefixation, and non-realization of infix. The evidence indicates that avoidance of these sequences applies only within the root-and-infix domain, and only in derived environments. This diversity of repairs seems unexpected if changes should be perceptually minimal; we suggest possible explanations.

Keywords Infixation · Austronesian · Obligatory Contour Principle · Dissimilation, heterogeneity of process · Too-many-solutions problem · Phonological typology

1 Diverse repairs for multiple labial consonants

The relationship between constraints on surface forms and operations that alter representations is of central interest in phonological theory. Classic Optimality Theory (OT; Prince and Smolensky 1993/2004) predicts that, cross-linguistically, multifarious “repairs” can occur in response to a markedness constraint. For example, Pater (1999—though see Pater 2001) describes diverse ways in which languages avoid a

K. Zuraw (✉)

Department of Linguistics, University of California, Los Angeles, 3125 Campbell Hall, Box 951543,
Los Angeles, CA 90095-1543, USA
e-mail: kie@ucla.edu

Y.-A. Lu (✉)

Department of Linguistics, Stony Brook University, Stony Brook, NY 11794-4376, USA
e-mail: yualu@ic.sunysb.edu

sequence of nasal consonant plus voiceless obstruent; Kennedy (2003) shows that Micronesian languages have made varied repairs to initial geminate consonants. For two recent discussions, see Baković (2007) on feature agreement and Yip (2007) on tonal OCP. This “heterogeneity of process” (McCarthy 2002) was the motivation for Kisseberth’s (1970) seminal proposal to incorporate explicit surface constraints into phonological theory. On the other hand, Steriade (2001/2008) points out that processes are sometimes not so heterogeneous. For example, a cross-linguistically widespread constraint against (word- or syllable-) final voiced obstruents seems to always trigger the same repair, final devoicing (Lombardi 1995/2001). Classic OT over-predicts the typology: we would also expect deletion, manner change, metathesis, and epenthesis as repairs for a final voiced obstruent. This “too-many-solutions” problem (Steriade’s term) has been taken up by Wilson (2001), De Lacy (2003—cited in Blumenfeld 2006), Pater (2003), Blevins (2004), Blumenfeld (2006), van Oostendorp and Hermans (2007), and others.

It remains unclear which constraints elicit a variety of repairs and which are consistent in the repair they trigger. The purpose of this squib is to add to our typological knowledge another case of diverse solutions. Like Kennedy’s, this case involves a cognate structure in related languages, allowing a relatively direct comparison.

The Proto-Austronesian infix **-um-* (Ross 2002; Blust 2003a) has descendants in many modern Austronesian languages.¹ The phenomenon of interest is how this infix behaves in stems beginning with (or, in some cases, merely containing) labial consonants. If stem and infix surface unchanged, a violation occurs of the cross-linguistically widespread prohibition on homorganic consonants within a stem, as in Tagalog *p-um-ili* ‘choose’ (OCP [Obligatory Contour Principle] as applied to segmental features: McCarthy 1979; Mester 1986; Yip 1989; Frisch et al. 2004, and many others). In some of the cases below, the same “repairs” that occur with labial-initial stems occur with stems beginning in nasals, suggesting avoidance of successive nasal consonants. A preliminary analysis of *-um-* among such languages is discussed by Lu (2005, in progress).

We review strategies for dealing with *labial...labial* sequences: tolerance (Sect. 2), gaps (Sects. 2 and 8), loss of stem consonant (Sect. 3), loss of infix nasality (Sect. 4), stem dissimilation (Sect. 5), infix dissimilation (Sect. 6), prefixation (Sect. 7), and non-realization of infix (Sect. 8). The existence of a prefixation strategy suggests that the constraint on *labial...labial* sequences applies to some domain that includes the stem and infix, but not a prefix (Łubowicz to appear); other evidence for the exclusion of prefixes is discussed below.

We also provide a very compressed sketch of an analysis for each language. Throughout there is a tension between two basic analytic approaches: (i) infixation is the norm but a markedness constraint prevents it in labial-initial stems, or (ii) something other than infixation is the norm, but faithfulness constraints prevent that other strategy in all but the labial-initial stems. We will point out cases in which

¹We are not aware of research directly addressing the question of whether the proto-infix was **um* or **m*, since both are found in modern languages. Presumably, the authors we cite choose **um* because the modern languages that have a vowel-*m* infix use a reflex of **u* instead of some other vowel. It is not implausible that in the vicinity of *m* the favored epenthetic vowel would be *u*, but perhaps the uniformity of the choice is viewed as unlikely to occur by that means alone.

both approaches work (and one where approach (ii) is needed), but in general the markedness-driven approach (i) is more successful. We will conclude that most of the languages can be analyzed as having a high-ranking markedness constraint *[_{stem}LAB V₀ LAB that rules out a stem-initial labial when the next consonant is also labial (as in *p-um-ili*), though a few languages display a more general or more specific version of an OCP-labial constraint, and some additionally obey OCP-nasal or OCP-sonorant.

To facilitate comparison across languages, transcriptions here are converted to IPA according to the source's description. For example, if the source indicates that words it transcribes as vowel-initial actually begin with [ʔ], the [ʔ] is shown here. However, some fine details of segmental allophony described but not transcribed in the source are ignored. The “stem” given is sometimes inferred from its affixed form(s). Data are cited by abbreviated source and page number, unless the source is a dictionary, in which case page number is usually omitted.

2 Tolerance and gaps

In Tagalog (data from English 1986 [E]), the realization of the infix is fairly uniform, as shown in (1)—cells for stems with labial consonants are outlined—although there are gaps, discussed below.

(1) Tagalog /-um-/

<i>stem-initial C</i>	<i>stem</i>	<i>actor focus, infinitive/past</i>	<i>gloss</i>
p	pili	p-um-ili	‘choose’
t	takot	t-um-akot	‘frighten’
k	kanta	k-um-anta	‘sing’
b	bukas	b-um-ukas	‘open’
d	damaj	d-um-amaj	‘sympathize’
g	gawaʔ	g-um-awaʔ	‘make’
s	sulat	s-um-ulat	‘write’
m	gap		
n	nipis	n-um-ipis	‘become thin’
ŋ	ŋitiʔ	ŋ-um-itiʔ	‘smile’
l	lipat	l-um-ipat	‘move’
r (loans only)	radjo	r-um-adjo	‘radio’
ʔ	ʔawit	ʔ-um-awit	‘sing’
h	hiŋiʔ	h-um-iŋiʔ	‘ask for’
w	partial gap		
	wagajwaj	w-um-agajwaj	‘wave’
j	jakap	j-um-akap	‘embrace’

Schachter and Otones (1972) note a lack of *m*- or *w*-initial bases taking *-um-*, though English does list two *w-um-* cases. There are few *m*- and *w*-initial stems in Tagalog, and not all verbs take the *-um-* infix, so one might wonder if this is an accidental gap. Orgun and Sprouse show that the gap is real by documenting consultants’ acceptance of nonce loans with *-um-* (*p-um-ejnt* ‘paint’) except when the stem begins in *m* or *w*: **m-um-eri* for ‘marry’, **w-um-ejl* for ‘wail’.

An analysis of these facts must appeal to the markedness of **m-um-eri*, as it is implausible that the basic strategy is to have a paradigm gap, with gaps somehow forbidden for stems beginning in anything but a labial sonorant. For example, modifying Prince and Smolensky's 1993/2004 constraint MPARSE (forbidding paradigm gaps) to MPARSE_[stem] {-LAB, -SON} seems absurd. Thus, we are forced to take the markedness approach. The most general markedness constraint that could be formulated is **_[stem]SONLAB V₀ LAB*, which forbids a stem from beginning with a sonorant labial if the next consonant is also labial. By design, the constraint does not rule out *m...m* sequences produced by prefixation. Like many of the languages discussed here, Tagalog has a prefix *ma-* forming verbs and adjectives, but there are no apparent restrictions on its attachment to labial-initial stems: *mulat* 'open (of eyes)' *ma-mulat* 'have one's eyes opened' (E); *wala?* 'absent' *ma-wala?* 'be lost' (E). Thus the constraint's domain should be restricted to the stem, as suggested by Łubowicz (to appear—see discussion of Palauan below), which includes the infix.² Nor does the presence of a labial later in the stem pose any problem for the constraint: *d-um-ama-j*. This markedness constraint outranks MPARSE (with perhaps some variation, to account for the possibility of *w-um-agajwaj*); the faithfulness constraints that would be violated by other solutions also outrank MPARSE.

The restriction of this OCP constraint to stem-initial position echoes the finding of Frisch (1996) that in Arabic, "OCP-Place is more strictly enforced at the beginning of the word" (p. 109). Frisch offers an explanation in terms of phonological encoding. Under the model he adopts there is little interference from other lexical items as a word-initial consonant is being accessed, so its similarity to the next consonant can be robustly assessed. Later in the word, competition from other lexical items interferes with this assessment. Frisch (2004) also finds that similarity-based consonant co-occurrence restrictions in English are strongest in word-initial syllables (and stronger in second than in third syllables).

3 Loss of stem consonant (analyzed as fusion)

There are several languages in which labial-initial stems begin with *m* instead of the expected *C-um-*.³ The pattern is shown in (2) for Timugon Murut (Prentice 1971:

²Monomorphemic stems largely obey the OCP: in Tagalog it is very rare to find two labial consonants within an unreduplicated root, as presumably in the other languages here too. But, there is a systematic exception for pseudo-reduplicated roots, which take the form *C₁V₁C₁V₁(C₂)* or *C₁V₁C₂C₁V₁C₂* (with optional pseudo-infixes such as the *ag* in *wagajwaj*), such as Orgun and Sprouse's example *mumo?* 'ghost (children's term)', where [o] is the final-syllable allophone of /u/. Perhaps the pseudo-reduplicant (*mu-*, in the case of *mumo?*) does not count as part of the stem. See Zuraw (2002) for other ways in which Tagalog pseudo-reduplicated roots act as though they were morphologically complex.

Loans are also apparently exempt (e.g., *mamon* 'sponge cake' < Spanish), though it is not known whether OCP-violating words are borrowed as freely as others. This special behavior of loans could be accounted for by assigning loans to a different component of the lexicon (Itô and Mester 1995a, 1995b, 1999), or perhaps by making the realization of a root a stronger requirement than the realization of an *um*-infix form. (These two analyses are also available for the pseudo-reduplicated roots.)

³As Blust (2004) points out, the scattering of this phenomenon across the language family means that it was independently innovated in multiple languages, or it was independently wiped out by analogy in multiple languages (or both).

130–131), in (3) for Mayrinax Atayal (Li 1995 [L]; Huang 2000 [H]), in (4) for Sarangani Manobo (DuBois 1976 [D]), and in (5) for disyllabic and CC-initial stems in Sarangani Blaan (also known as Bilaan; Rhea 1972, 1995 [R₇₂, R₉₅] and Rhea's contributions to the comparative wordlist of Tryon 1995a [T—with a subscript indicating volume number]; Abrams 1961, 1970 [A₆₁, A₇₀]; McLachlin and Blackburn 1968 [MB]; unfortunately there is a mixture of systems for transcribing vowels). For discussion of infixation in Mayrinax Atayal, see Lu (2005).

(2) Timugon Murut /-um-/⁴

	<i>stem</i>	<i>future (subject focus, intransitive Vs)</i>	<i>gloss</i>
p	patoj	m -atoj	'die'
t	tuun	t-um-uun	'go down'
b	bigod	m -igod	'stand'
g	gajo	g-um-ajo	'grow bigger'
m	mitiŋ	m -itiŋ	'hold a meeting'
V	oŋoj	m -oŋoj	'go'
k, d, s, n, ŋ,	<i>no data</i>		
l, ʔ, w, j			

(3) Mayrinax Atayal /-um-/

	<i>stem</i>	<i>actor focus</i>	<i>gloss</i>
p	paqut	m -aqut	'ask' (L286)
t	timami?	t-um-imami?	'pickled meat' (L290)
c	curiŋ	c-um-uriŋ	'burn' (H388)
k	kai?	k-um-ai?	'dig' (L285)
q	qumi?	q-um-umi?	'pickle' (L290)
s ⁵	sauq	s-um-auq	'smell' (H388)
β	βuq	m -uq	'squeeze' (L286)
ɣ	ɣaliq	ɣ-um-aliq	'tear' (L286)
l	laɲuj	l-um-aɲuj	'swim' (L286)
r	rahiŋ	r-um-ahiŋ	'air-dry' (L286)
ʔ	ʔiŋat	ʔ-um-iŋat	'rob' (H387)
h	haβ	h-um-aβ	'stab' (L286)
x, m, n, ŋ, w, j	<i>no data</i>		

(4) Sarangani Manobo /-om-/

	<i>stem-initial C</i>	<i>stem</i>	<i>future</i>	<i>gloss</i>
p	paŋgas	m -aŋgas		'sow' (D18, 19)
t	toman	t-om-oman		'fulfill' (D18, 19)
k	kaən	k-om-aən		'eat' (D18, 19)
b	bajad	m -ajad		'buy' (D18, 19)

⁴Prentice states that in Timugon Murut there is one exceptional stem beginning with *d*: *datoŋ* vs. *m-atoŋ* 'come'.

⁵Huang has "ʃ", described as an alveo-palatal fricative, where Li has [s].

d	datəŋ	d-om-atəŋ	‘arrive’ (D18, 19)
g	gojod	g-om-ojod	‘pull’ (D18, 19)
s	samsam	s-om-amsam	‘hike at night’ (D19)
l	likoʔ	l-om-ikoʔ	‘return again’ (D19)
ʔ	ʔadek	m -adek	‘smell’ (D18, 19)
h, m, n, ŋ, w, j	<i>no data</i>		

- (5) Sarangani Blaani /-m-/; vowel transcription in each item is that of the first source cited

	<i>mono-</i>	<i>actor</i>	<i>disyll-</i>	<i>actor</i>	<i>gloss</i>
	<i>syll-</i>	<i>focus</i>	<i>abic or</i>	<i>focus</i>	
	<i>abic</i>		<i>CC-</i>		
	<i>stem</i>		<i>initial</i> ⁶		
			<i>stem</i>		
t	təb	m-təb			‘graze’ (R ₉₅ 372)
			təjud	t-m-əjud	‘endure’ (T ₄ 117)
			tbel	t-am-bel	‘sew’ (R ₉₅ 372)
k			kəʔən	k-m-əʔən	‘eat’ (T ₂ 629, A ₆₁ 396)
			klaŋ	k-am-laŋ	‘cut’ (A ₆₁ 399)
b	bat	m-bat			‘throw’ (A ₆₁ 399, MB119)
			bunal	m -unal	‘beat’ (R ₉₅ 372)
			blé	m -lé	‘give’ (A ₆₁ 398, MB119)
d			dado	d-m-ado	‘plow’ (R ₇₂ 38)
			dsuʔ	d-am-suʔ	‘sacrifice’ (R ₇₂ 39, MB119)
f (< *p)	fuk	m-fuk			‘wash clothes’ (T ₃ 335, MB119)
			fati	m -ati	‘kill’ (R ₉₅ 372)
			fwes	m -wes	‘uncover’ (MB119)
s			saloʔ	s-m-aloʔ	‘hunt’ (MB109)
			stifun	s-am-tifun	‘assemble’ (A ₇₀ 2)
l			lifo	m -lifo ⁷	‘cook’ (T ₂ 656, R ₇₂ 38)
ʔ			ʔimoʔ	m -imoʔ	‘make’ (MB119)
g, m, n, ŋ, r, h, w, j			<i>no data</i>		

Following Carter (2000) and Pater (2001), we take the stem consonant and the affixal /m/ to be fused, in violation of McCarthy and Prince’s (1995) UNIFORMITY; we further assume that the resulting [m] is not incorporated into the stem: /m₁ + p₂atoj/ → m_{1,2}-[_{stem}atoj]. Because only a labial consonant can fuse with /m/ without a violation of McCarthy and Prince’s (1995) IDENT constraints for various place features, it is possible to analyze these languages with fusion as the norm (perhaps to satisfy McCarthy and Prince’s 1993 ALIGN constraints), but blocked by IDENT for all but the labial consonants.

⁶If the stem is CC-initial, the infix is realized as -am-.

⁷None of the authors comment on *l*, and the only two examples are in Rhea (1972). Both take initial *m*-.

If we are to pursue a markedness-driven account, the markedness constraint involved must be more general than that introduced for Tagalog above, affecting obstruents and sonorants alike. However, the exclusion of prefixes still holds: Mayrinax Atayal, like Tagalog, has a *ma-* prefix that does not trigger any repair: *βicuw*, *ma-βicuw* ‘thunder’ [H387, but with Lu’s (2005) morphological analysis]. And, labials later in the stem still have no effect (e.g., Mayrinax Atayal *q-um-ami?*). We invoke the constraint mentioned in the introduction, *[_{stem}LAB V₀ LAB, which forbids a stem-initial labial when the next consonant is labial. This constraint outranks UNIFORMITY (and MAX-V, when there is an underlying vowel in the affix).⁸

Kapampangan (del Corro 1980 [dC]; Forman 1971 [F]) has a similar pattern, but only, according to del Corro,⁹ when the first vowel of the stem is short (6); fusion is optional, at least for some words.

(6) Kapampangan /-um-/

	<i>stem, Ist V long</i>	<i>actor focus, future</i>	<i>stem, Ist V short, Ist syll. open</i>	<i>actor focus, future</i>	<i>stem, Ist V short, Ist syll. closed</i>	<i>actor focus, future</i>	<i>gloss</i>
p	pa:la	p-um-a:la		pusit	m -usit	paldak	‘shovel’ (dC46)
						m -aldak	‘squeeze out’ (dC42)
						~paldak	‘kick’ (dC43,46)
t	te:rak	t-um-e:rak		talakad	ta:lakad	tikdo	‘dance’ (F105)
						tikdo	‘stand’ (dC46)
							‘stand’ (dC45)
b	ba:ge	b-um-a:ge		bulus	m -ulus		‘fit’ (dC46)
				buluŋ	m -uluŋ		‘pour’ (dC42)
					~bu:luŋ		‘whisper’ (dC42,46)
						baldug	‘drop’ (dC43)
						m -aldug	

⁸For Murut vowel-initial stems, we additionally need MAX-V_{root} » ONSET » MAX-V.

For Blaen, there are a number of other complexities in the grammar, including a constraint against tri-consonantal onsets (triggering epenthesis of *a*) and a ban on infixing into a monosyllabic base (outranking the affix’s preference to be an infix). To account for prefixation on *l*-initial stems, we need another constraint *[_{stem}SON V₀ SON prohibiting a stem-initial sonorant from being followed by another sonorant; fusion (/m₁/, /l₂ifo/ → *m_{1,2}-ifo) is ruled out by IDENT(place).

We are forced to treat initial glottal stops in Blaen and Manobo as epenthetic to account for their non-appearance in infixed forms (i.e., to rule out *ʔ-om-adek or *m-ʔadek in favor of *m-adek*); apparently, Manobo prefers to obey ONSET by deletion of the affix’s /o/ rather than insertion of ʔ. In Atayal, the ʔ could be epenthetic (if its insertion is preferred to vowel deletion) or underlying (if not).

⁹Some of Forman’s data contradict del Corro’s generalizations, but since del Corro goes through the patterns in more detail and more systematically, we have reported her claims here. Of course it is possible that the two authors are describing different varieties.

g	ga:to	g-um-a:to				'float' (dC46)
			<i>no data</i>			
				gasgas	gasgas	'rub against' (dC46)
s	sul:at	s-um-u:l:at				'write' (dC46)
			sal:i:lun	sa:li:lun		'go under shade'
				<i>no data</i>		(dC46)
n	na:nu	n-um-a:nu	<i>no data</i>	<i>no data</i>		'do' (F104)
l	li:nis	l-um-i:nis				'clean' (dC46)
			la:bug	la:bug		'become blurred'
						(dC45)
				lislis	lislis	'put up skirt' (dC46)
V ¹⁰	ʔi:sip	m-i:sip				'think' (dC39)
			ʔasuk	m-asuk		'emit smoke' (dC40)
				ʔugse	m-ugse	'throw' (dC40)
k, d, m, ŋ,			<i>no data</i>			
r, h, w, j						

We assume an arbitrary constraint in Kapampangan forbidding *um*, in its full form, from being followed by a short vowel; when the stem's first vowel is short, the /m/ is therefore lost and the /u/ fuses with the stem vowel, lengthening it (unless the syllable is closed; superheavy syllables are not allowed): /u₁m/, /t₂alakad/ → [ta:_{1,2}alakad]. (Vowel fusion is more apparent in past-tense forms, derived from adding the infix /in/ to the future-tense base: /i₁n/, /mu₂lus/ → [mi:_{1,2}lus] 'poured' [dC42].) If the stem-initial consonant is labial, however, vowel fusion and loss of /m/ is optionally avoided by instead fusing the consonants (and losing the vowel). If this is the right analysis, then an OCP markedness constraint is not actually the driving force behind labials' special behavior—rather, fusion is (optionally) preferred over deletion for avoiding [um]-plus-short-vowel, but it is available only when the stem consonant agrees with the affix's /m/ in underlying place.¹¹

In Muna (van den Berg 1989: 32–35; analyses in Carter 2000; Pater 2001), *p*, *f*, and sometimes *w* are replaced by *m*, but *b* and *ɸ* simply show no reflex of the infix. Stems that are *m*-initial are ambiguous between stem-consonant loss and no reflex. Nasal-initial, prenasalized-C-initial (including ^m*b*), and some *w*-initial stems show no reflex:

¹⁰In certain *ʔaCC*-initial stems, the initial *a* is lost, the infix is inserted between the two consonants, and the nasal of the infix assimilates to the following consonant: *ʔaslam*, *s-un-lam* 'become sour' [dC, 38]. Del Corro does not indicate whether words transcribed with initial vowels actually begin with a glottal stop, but in the recording accompanying Balilo (2008), it is clear that they all do, at least in citation form. To account for these words' /m/ forms, we must assume that the glottal stop is epenthetic.

¹¹In more detail: IDENT(place) » MAX-C, UNIFORMITY-V which is why /m/ loss and vowel fusion are used for non-labial-initial stems; but MAX-C and UNIFORMITY-C are variably ranked, so both that option and consonant fusion are available in labial-initial stems. A markedness-based analysis is unsupported, because **b-um-uluy* is already ruled out by the arbitrary constraint *UM + SHORT-V, leaving no role for an OCP constraint (although it would be possible to rank one fairly high). Because both *bu:luy* and *m-uluy* satisfy any OCP-labial constraint, it must be the difference in faithfulness that distinguishes the two.

(7) Muna /-um-/

	<i>stem</i>	<i>irrealis</i>	<i>gloss</i>
p (<*p ¹²)	pili	m -ili	‘choose’
b (<*b)	baru	baru (<i>zero</i>)	‘happy’
ɓ (<*b)	ɓala	ɓala (<i>zero</i>)	‘big’
d _n	dudu	d _n -um-udu	‘push’
d	dadi	d-um-adi	‘live’
g	gaa	g-um-aa	‘marry’
f (<*p)	foni	m -oni	‘climb, go up’
s	solo	s-um-olo	‘flow’
^m b	^m bolaku	^m bolaku (<i>zero</i>)	‘steal’
ⁿ d	ⁿ diwawa	ⁿ diwawa (<i>zero</i>)	‘yawn’
m (<*m)	manda	m -anda (<i>zero or m</i> -)	‘repent’
n (<*n, ñ, ŋ)	nale	nale (<i>zero</i>)	‘soft, weak’
l	limba ¹³	l-im-imba	‘go out’
r	rende	r-um-ende	‘alight’
h	hela	h-um-ela	‘sail’
w ¹⁴ (<*b)	waa	m -aa	‘give’
	wanu	wanu (<i>zero</i>)	‘get up’
V	ala	m -ala	‘take’
t, k, ɸ, ^ɰ g, ⁿ s, ŋ	<i>no data</i>		

Carter analyzes Muna stem-consonant deletion as fusion, driven not by the OCP but by ALIGN constraints (McCarthy and Prince 1993) that prefer both the affix and the stem to begin at the left edge of the word. It is only when the stem-initial consonant’s features are incompatible with fusion that infixation occurs, and the cases of zero realization are, for Carter, coalescence in which the affix’s unique features do not survive. As noted above, this alignment analysis is available for other cases of fusion. If we do adopt a markedness-driven analysis, then both *[_{stem}LAB V₀ LAB and *[_{stem}NAS V₀ NAS outrank UNIFORMITY.¹⁵

Palauan (Flora 1974; Josephs 1990—data below all from Josephs 1990) also shows loss of stem-initial /b/ (8). But in addition, when a labial consonant appears *later* in the stem, the stem consonants remain unchanged and the infix changes from its usual [m] to [w]—strangely, since [w] is also labial and we have seen that it participates in OCP-labial in Tagalog and Muna (diachronic developments seem to offer no insight here, as modern [w] can come from **u/w* or **p*, all labial). Either the prohibition in Palauan is on successive labial, [+consonantal] segments, or [w], although non-syllabic on the surface, is treated as an allophone of /u/ and thus exempt from

¹²Proto-Malayo-Polynesian source, as reconstructed by van den Berg (1991).

¹³The infix appears as *-im-* because the following vowel is *i*.

¹⁴“Some of the [*w*-initial] roots undergo nasal substitution, but the majority do not change” (p. 33).

¹⁵And, once again, ONSET forces deletion of the affix’s /u/ in vowel-initial stems.

OCP considerations. The dissimilation seems to be obligatory when the very next consonant is labial, and optional (at least in one word) when the labial is later.¹⁶

(8) Palauan /-m-/¹⁷

	<i>stem</i>	<i>present perfective, plural non-human object</i> ¹⁸	<i>gloss</i>
t	toŋakl	t-m-oŋakl	‘put up high’
k	kiwt	k-m-iwt	‘clean up’
b	basəʔ	m -asəʔ	‘name’
s	sisijʔ	s-m-isijʔ	‘strengthen’
ð	ðakl	ð-m-akl	‘bury’
m	<i>no clear cases</i>		
ŋ	ŋas	ŋ-m-as	‘subtract’
l	latk	l-m-atk	‘remember’
r	roðəʔ	r-əm-oðəʔ	‘aim at blindly’
ʔ	ʔaðəs	ʔ-əm-aðəs	‘pave’
some w/u/o (possibly those that < *p)	oak	m -ak	‘anchor’
other V (few cases)	uʔət	im -uʔət	‘unhook’
	asəm	im -asəm	‘try out’
next C is labial	ðobəʔ	ð- w -obəʔ	‘chop down’
	kɛməð	k- w -ɛməð	‘sew up’
	jɛbəð	i- w -ɛbəð	‘lasso’
labial C later in word	ðaləm	ð- w -aləm	‘plant’
	sɛsəb	s- w -ɛsəb ~ s-m-ɛsəb	‘burn’

Palauan provides further evidence for a prefix/infix asymmetry, as discussed by Łubowicz (to appear). Previous authors have analyzed Palauan -m- and mə- as allomorphs of the same morpheme, whose distribution is determined largely by the verb’s morphosyntactic features.¹⁹ Josephs (1975) calls this morpheme the “verb marker”; it appears as a prefix [mə-] or [o-] in ergative verbs and “many” (p. 146) intransitive action or state verbs (mə-rael ‘walk’). If the stem begins in a labial, the prefix must appear instead as o-: o-boes ‘get shot’; Josephs characterizes this as dissimilation. (There are also “a small number of verbs” (p. 147) that take o- despite having no labial consonant.) For most intransitive verbs, and all perfective verbs, the verb marker shows up instead as the infix -m- (l-m-uwt ‘return_V’, cf. luwt ‘return_N’), with the allomorph -w-²⁰ (or -o- if certain additional phonology applies) as noted

¹⁶Personal communication from Andrew Martin, who compiled a list of infixed forms of meN- prefixed stems in Josephs that have a non-initial labial.

¹⁷Palauan has no [p]. Earlier *p became [w], [u], or sometimes [o].

¹⁸This affix occurs in a variety of verb forms, but these particular perfectives afford the best opportunity to observe it unobscured by additional phonological changes.

¹⁹One could speculate that proto-Austronesian *ma- and *-um- have merged.

²⁰Notated u by Josephs—see his Chap. 1.

above. Łubowicz makes the interesting point (see also Finer 1986) that for the prefixal forms, dissimilation of *mə-* to *o-* is local on the consonantal level—it applies only when the stem-initial consonant is labial (*o-boəs* instead of **mə-boəs*). But for the infixal forms, dissimilation can be long-distance. (See Lu (in progress) for analysis of long-distance dissimilation in Palauan and local dissimilation in Mayrinax Atayal and many others.) Łubowicz attributes the difference between prefixes and suffixes to their morphological affiliation—specifically, she proposes that infixes reaffiliate to the root and are then subject to stricter consonant cooccurrence restrictions. The data above offer further support for this view: vowel-initial stems mostly take prefixal *im-*, and in the one example found where there was a labial consonant later in the stem, no dissimilation occurs: *im-asəm* ‘try out’. The *m-* allomorph seen in *b-*initial perfectives does not undergo further dissimilation (from root *bəob*, *m-əob* ‘form’ instead of **w-əob*); if we wish to maintain Łubowicz’s view of morphological affiliation, we must analyze the [m] in *m-əob* as prefixal, whether it results from fusion of /m/ and /b/ or deletion of the /b/.

We analyze Palaun /b/-fusion as driven by $*[\text{stemLAB } V_0 \text{ LAB} \gg \text{UNIFORMITY}]$ (as in Muna), and dissimilation of /m/ to [w] as driven by $*[\text{stemLAB } V_0 \text{ LAB} \gg \text{IDENT(continuant)}]$ —accounting for the exceptionless dissimilation when the next consonant is labial—plus a strong tendency for a more general markedness constraint, $*[\text{stemLAB} \dots \text{LAB}]$, to outrank IDENT(continuant) —accounting for the strong tendency towards dissimilation when any following consonant is labial.²¹ The dissimilation of the other affix, /mə-/ to *o-* before stem-initial /b/ is driven by a constraint that is local but not confined to the stem domain: $\text{LAB } V_0 \text{ LAB}$.

Aborlan Tagbanwa is another language that shows deletion of stem-initial labials (Green 1979),²² though only one example is available: the Actor-Focus form of stem *betang* is *metang* (p. 70—transcriptions as in the original). Some other examples are Kimaragang (*p* only; Kroeger 1988: 222), and Sediq (a.k.a. Taroko: Li 1991; Yang 1976).

4 Loss of infix nasality

We have found one case in which the nasality of the infix /m/ of the infix is lost entirely. The language is Kavalan (Li and Tsuchida 2006). As shown in (9), /-m-/ shows up as an infix with voiceless stem-initial consonants (stop and fricative) and as a prefix with voiced stem-initial consonants (fricatives, nasals, liquids and glide). However, the affix appears as [u] before a stem-initial voiced labial fricative /β/.

²¹There are several complications. An ALIGN constraint prefers for the word to begin with material from the underlying root (forcing infixation, since word-initial *mC* is not in general ruled out phonotactically). This outranks a constraint penalizing re-affiliation of affixal material to the stem, but that constraint has an emergent effect in forms like *m-asəʔ*, ensuring that the initial [m] does not count as part of the stem (we assume that ALIGN is satisfied in these cases, because the initial [m] still corresponds to the underlying root-initial /b/, as well as to the affixal /m/). The *im-* allomorph seen with vowel-initial roots is puzzling; it could be analyzed with a correspondence constraint requiring the /m/-affixed forms of vowel-initial roots to remain vowel-initial, though this seems typologically unexpected.

²²“In a stem beginning with a bilabial consonant -um- manifests as m- and portmanteau occurs in which the bilabial is lost.” (p. 88)

(9) Kavalan /-(u)m-/²³

<i>stem-initial C</i>	<i>stem</i>	<i>actor focus</i>	<i>gloss</i>
p	paʔit	p-m-aʔit	'tie'
t	tajta	t-m-aʔta	'see'
k	kirim	k-m-irim	'seek'
q	qan	q-m-an	'eat'
β	βura	u-βura	'give'
s	sum	s-m-um	'urinate'
z _i	zap	m-zap	'find'
m	<i>no data</i>		
n	nanap	m-nanap	'crawl'
ŋ	ŋaβ	m-ŋaβ	'open the mouth'
l	lames	m-lames	'clean'
r	rinaw	m-rinaw	'do dishes'
ʋ	ʋapis	m-ʋapis	'split'
w	warin	m-warin	'throw away'
V	ara	m-ara	'take'

It is unclear whether the underlying form should be /um/ or /m/, and thus it is unclear whether we are seeing deletion of /m/ before /β/ (with /u/ deleting elsewhere), or merely its denasalization. One possible analysis appeals to sonority sequencing to drive the basic infixation/prefixation split: [m] can occur after stops and voiceless fricatives, but not after voiced fricatives, nasals, liquids and glides (SONORITYSEQUENCING » ALIGN).²⁴ With initial /β/, however, simple prefixation would create the sequence a labial-labial sequence, albeit across the prefix-stem boundary [... m-β. ...]. We have seen that in Palauan, such a configuration is subject to dissimilation (*mə-* becomes *o-* before *b*).

This pattern can be analyzed with a constraint *LAB VOICEDLAB, banning two adjacent labial consonants when the second is voiced. (As in Palauan, vowels and the glide [w] are exempt.) This constraint is like the Palauan constraint on the prefix /m-/ in that it is local but not confined to the stem. This markedness constraint outranks MAX (if /u₁m₂, b₃.../ → [u₁-b₃...]) or IDENT(nasal) and IDENT(consonantal) (if /m₁, b₂.../ → [u₁-b₂...]).²⁵ Infixation of [u] or [w] is also ruled out, perhaps by general phonotactic constraints: *[β-w-ura], *[β-u-ura].

²³In stems that begin with a consonant cluster, the /u/ appears either before or after /m/. When the first consonant is /p, t, k, q, s/, -*um-* appears as an infix: *p-um-laq* 'squeeze'. When the first consonant is /l, r, ʋ, z, n/, *mu-* appears as a prefix: *mu-lnap* 'whisper'.

²⁴This requirement would apply only in derived environments, since /mp, mt, mk, mq, ms/ and /zm, lm, rm, ʋm/ are all attested word-initial consonant clusters (Li and Tsuchida 2006: 4).

²⁵Although *LAB VOICEDLAB outranks MAX/IDENT, MAX/IDENT outranks ALIGN, so that prefixation rather than a change in sonority is used to avoid a sonority sequencing violation for the rest of the voiced consonants.

5 Stem dissimilation

There are also languages in which a labial stem-initial consonant dissimilates in place. As shown in (10), in Limos Kalinga (Ferreirinho 1993 [F]; Wiens 1979 [W]) the stem consonant becomes velar.²⁶ We analyze this as *[_{stem}LAB V₀ LAB » IDENT(place)/stem. Only a markedness-driven analysis is plausible—it seems unlikely that there would be a blanket preference for changing the stem consonant to a velar, but that this is not possible for coronals.

(10) Limos Kalinga /-um-/

	<i>stem</i>	<i>actor focus, inchoative</i>	<i>gloss</i>
p	pija	k -um-ija	‘good’ (F25)
k	kan	k-um-an	‘eat’, ‘bite’ (W24, F41)
b	bulbul	g -um-ulbul	‘cook rice’ (W25, F25)
d	datoŋ	d-um-atoŋ	‘come’ (F24)
l	lamʔok	l-um-amʔok	‘soft’ (F25)
ʔ	ʔudan	ʔ-um-udan	‘rain’ (F24)
t, g, s, m, n, ŋ, h, w, j	<i>no data</i>		

As shown in (11), in Northern Acehnese (Durie 1985 [D]), a stem-initial /p/ becomes a coronal fricative [S];²⁷ this happens only with disyllabic stems, however. For trisyllabic stems with any initial consonant (not shown in (11) except for /p/), the morpheme shows up as prefix [mu-] (or [mu-]: when not bled by the /p/ → [S] change, /u/ → [u] between labial consonants) instead. The prefixal allomorph also appears before stems—regardless of length—beginning in /b,ʔ,h,l,r,m,n,p,ŋ,w,j/, suggesting avoidance not only of *labial* ... *labial* but also *sonorant* ... *sonorant* within the stem domain (assuming that glottals count as sonorant). Monosyllabic stems (not shown except for /p/) are augmented before being infixes, either by epenthesis into an initial consonant cluster, or by reduplication of the initial consonant followed by [u].²⁸ Or, for “some” (p. 96) monosyllabic *p*-initial roots, the /mu-/ prefix is used ([plɛ], [mu-plɛ] ‘pour’). We omit spreading of nasality from the examples.

(11) Northern Acehnese /-uum-/

	<i>stem</i>	<i>intransitive</i>	<i>gloss</i>
p	<i>disyllabic:</i> pubu ⁹ t	S -uum-ubu ⁹ t	‘teach’ (D95)
	<i>trisyllabic:</i> puuraŋuj	mu -puuraŋuj	‘act’ (D94)
	<i>monosyllabic:</i> plah	S -uum-ulah	‘split’ (D95)
		~ S -uum-uplah	
t	tulak	t-uum-ulak	‘push’ (D95)
c	caŋkoj	c-uum-aŋkoj	‘hoe’ (D95)

²⁶Ferreirinho also gives an example that appears to be an exception: *bilog*, *b-um-ilog* ‘strong’ (p. 25).

²⁷Durie describes [S], “a sound for which there is no satisfactory IPA symbol”, as “a laminal alveo-dental fricative with a wide channel area” (p. 12).

²⁸Exception: if the monosyllabic root begins with *rh* [< *Sh] or *lh* [< *Sl], [Sui] is added to the beginning of the root.

k	kalɤn	k-um-alɤn	‘look’ (D95)
b	blɔ ^ə	mu -blɔ ^ə	‘buy’ (D94)
d	duɤɤ	d-um-ɤɤ	‘hear’ (D95)
ʃ	ʃarom	ʃ-um-arom	‘inject’ (D95)
g	gantoŋ	g-um-antoŋ	‘hang’ (D95)
S	Salu ^ə n	S-um-alu ^ə n	‘be a cry-baby’ (D95)
ʃ	<i>no data</i>		
m	mat	mu -mat	‘hold’ (D94)
n	nari	mu -nari	‘dance’ (D95)
ɲ	ɲapɔ ^ə	mu -ɲapɔ ^ə	‘sing’ (D95)
ŋ	ŋiɛŋ	mu -ŋiɛŋ	‘see’ (D95)
l	laɲu ^ə	mu -laɲu ^ə	‘swim’ (D94)
r	raba	mu -raba	‘feel’ (D95)
ʔ	ʔaŋkot	mu -ʔaŋkot	‘transport’ (D94)
h	hila	mu -hila	‘drag’ (D94)
w	wanwan	mu -wanwan	‘squeeze’ (D95)
j	ju ^ə	mu -ju ^ə	‘order’ (D95)

Acehnese can be analyzed, like Limos Kalinga, as having $*[\text{stemLAB } V_0 \text{ LAB} \gg \text{IDENT}(\text{place})/\text{stem}]$. The change to [S] is not available for /b/, however, suggesting that $\text{IDENT}(\text{voice})$, $*[\text{stemLAB } V_0 \text{ LAB} \gg \text{ALIGN} (\gg \text{IDENT}(\text{place})/\text{stem})]$, where ALIGN is a constraint, as in Palauan, requiring stem material to be initial and thus disfavoring prefixation. We assume that there is no voiced, non-labial segment available in the phoneme inventory that /b/ could change to without violating some other high-ranked faithfulness constraint. Similarly, the sonorants are unable to change to [S] because of $\text{IDENT}(\text{sonorant})$, and so prefixation results, because of the ranking $*[\text{stemSON } V_0 \text{ SON} \gg \text{ALIGN}]$. There are also constraints against infixing into a monosyllable (as in Sarangani Blaauw) or a trisyllabic stem. As with Limos Kalinga, a markedness-driven analysis seems necessary: it is implausible that change to [S] is the default behavior. It also seems implausible that prefixation is the default behavior, since it is not clear what would rule out prefixation for obstruents besides /b/.²⁹

6 Infix dissimilation

There are languages in which the infix dissimilates when the stem-initial consonant is labial. In Kulalao Paiwan (Ferrell 1982 [F]), illustrated in (12), the infix’s [m]

²⁹There are some interesting complications in the Northern Acehnese case (Durie 1985: 33–34). First, the /p/ → [S] change also optionally takes place when the causative prefix /pu-/- attaches to labial-initial disyllabic, verbal (not nominal) roots: /pu-mu?um/ → [Su-mu?um] ‘to amuse, entertain’ and /pu-baro/ → [Su-baro] ‘to renew’ (p. 33)—this suggests a restriction against *labial...labial* even in the prefix + stem domain. Second, stem-initial /p,b,m/ all optionally become [S] when the nominalizing infix /-um-/ is added, if the next stem consonant is also labial (Durie’s examples appear to be reduplicated): /p-um-ɤpɔh/ → [S-um-ɤpɔh] ‘beating’, /b-um-ɤblɔ^ə/ → [S-um-ɤblɔ^ə] ‘thing bought’, /m-um-ɤme/ → [S-um-ɤme] ‘thing brought’. And third, /t/ can also optionally become [S] in the environment $_\{l,r,n\}VC(C)V(C)\#$, “where the final syllable is not /tV(C)/”.

becomes [n]. (If the first vowel of the stem is a schwa and the following consonant will be identical to the infix nasal, then the infix is optionally inserted after the schwa, so that a *məm* or *nən* sequence becomes *əmm*, *ənn*.) The pattern is similar in Tjuabar Paiwan (Ho 1978, 1995 [H₇₈, H₉₅]), shown in (13), except that, as in Palauan, there is also a change when a labial consonant comes later in the stem. Ho states that when “the second syllable of the stem has a labial initial that is preceded by /ə/,” the infix’s [m] dissimilates to [n]. For discussion of infix dissimilation in Paiwan, see Tseng (2003).

(12) Kulalao Paiwan /-m-/

	<i>stem</i>	<i>actor focus</i>	<i>gloss</i>
p	piliq	p- n -iliq	‘choose’ (F10)
t	tulək	t-m-ulək	‘point’ (F286)
c	cani	c-m-ani	‘drop’ (F291)
ts (“post-palatal” F7)	tsauv	ts-m-auv	‘cover’ (F310)
k	kan	k-m-an	‘eat’ (F10)
q	quma	q-m-uma	‘construct fields’ (F11)
b	burəs	b- n -urəs	‘use water-squirter’ (F66)
d	dət	d-m-ət	‘go near’ (F68)
ɖ (“post-palatal” F7)	ɖail	ɖ-m-ail	‘(monkey) raid crops’ (F11)
g	gudəm	g-m-udəm	‘finish’ (F96)
s	sapuj	s-m-apuj	‘use fire’ (F11)
v	vuɭuq	v- n -uɭuq	‘throw’ (F353)
z	zacak	z-m-acak	‘arrange’ (F360)
m	matsa	m- n -atsa	‘see w. own eyes’ (F173)
ɭ	ɭaɭiak	ɭ-m-aɭiak	‘scatter’ (F132)
l (“post-palatal” F7)	lima	l-m-ima	‘use hands’ (F162)
r	ratrat	r-m-atrat	‘expose seed’ (F242)
First V = [ə],	ɭəmət	ɭ-m-əmət ~ ɭə-m-mət	‘despise’ (F11)
second C = infix C	pənəc	p- n -ənəc ~ pə- n -nəc	‘decide’
j, ʔ, w, j, V	no data		

(13) Tjuabar Paiwan /-əm-/

	<i>stem</i>	<i>actor focus</i>	<i>gloss</i>
p	pajsu	p-ə n -ajsu	‘pound (rice)’ (H ₉₅ 310)
t	təkəɭ	t-əm-(ə)kəɭ	‘drink’ (H ₇₈ 652)
c	culu	c-əm-ulu	‘bake’ (H ₇₈ 648)
k	kan	k-əm-an	‘eat’ (H ₉₅ 310)
q	qaqiv	q-əm-aqiv	‘call’ (H ₇₈ 649)
d	dapəs	d-əm-apəs	‘blow (with breath)’ (H ₇₈ 649)
ɖ	ɖukuɲ	d-əm-ukuɲ	‘bend’ (H ₇₈ 649)
g	giriɲ	g-əm-iriɲ	‘growl’ (H ₇₈ 661)
s	supu	s-əm-upu	‘count’ (H ₉₅ 310)
v	vəɭi	v-ə n -əɭi	‘buy’ (H ₇₈ 649)
z	zuruɲ	z-əm-uruɲ	‘push’ (H ₇₈ 654)

l	ləŋləŋ	l-əm-əŋləŋ	‘see’ (H78655)
l	ləvəc	l-əm-əvəc	‘scoop out’ (H78662)
r	rakac	r-əm-akac	‘pull’ (H78654)
later C is labial	təvə a	t-ə n -(ə)və a	‘answer’ (H95310)
b, ʃ, ts, m, n, ŋ, ʎ	no data		

Ho (1978) gives a long wordlist for five dialects of Paiwan: Butanglu, Tjuabar, Tjavuali, “Paiwan”, and Stimul. A similar pattern is seen in all the dialects, with [-ən-] for stems beginning in labial obstruents; Ho (1977: p. 615) also gives an *m*-initial example, *mulji*, *m-ən-ulji* ‘force’, for Butanglu (*m*-initial examples are not available for the other dialects). The wordlist shows some variation for the case in which the infix is followed by (ə)v. Butanglu and Tjavuali consistently have [ən], but Tjuabar, “Paiwan”, and Stimul have one or two exceptions,³⁰ such as Tjuabar *q-əm-əvuc* ‘extinguish’, vs. Butanglu *q-ən-əvuc* (p. 663). The Paiwan languages show no dissimilation with the prefix *ma-* (Lu 2005: 71).

The Kulalao Paiwan pattern can be analyzed with the ranking *[_{stem}LAB V₀ LAB » IDENT(place)/affix. In Tjuabar Paiwan, to account for the dissimilation triggered by a later stem labial (but only across schwa), we need a markedness constraint like *LAB(ə)LAB (perhaps restricted to within the stem), which also outranks Ident(place)/affix. Given that the infix in Tjuabar Paiwan has a schwa infix, this markedness constraint could also account for the basic dissimilation pattern there. As with stem-consonant dissimilation, a markedness constraint seems necessary for these languages: it is unlikely that the default infix has /n/ but changes it to [m] when the stem begins with a non-labial (and, in Tjuabar, the infix is not followed by (ə)-labial).

7 Prefixation

In some languages, /-um-/ is an infix in non-labial-initial stems but a prefix in labial-initial stems.³¹ This is the case in Toba Batak as described by Nababan (1981 [N]), the published version of a 1958 thesis and a 1965 dissertation. We will use /um/₁ to refer to the version of the affix that forms comparative adjectives and completive forms of transitive verbs. Before labials and nasals /-um-/₁ becomes a prefix and its /m/ assimilates in place to the stem-initial consonant (14). According to van der Tuuk (1867/1971 [T]) and Nababan, the other version of the affix, /um/₂, which forms intransitives, displays instead the loss-of-stem-consonant pattern of Sect. 3.

³⁰This may be a coincidence, but in all the exceptions with [əm], a schwa intervenes between infix and *v*, and in the [ən] cases, the *n* and *v* are immediately adjacent (except in Butanglu, where an optional schwa is shown for some cases, that nonetheless have [ən]).

³¹Prefixation not driven by OCP-labial occurs in Chamorro, where, at least for some speakers, sonorant-initial stems take prefixation of *mu-* instead of the usual *-um-* infixation, and *ni-* instead of the usual *-in-*. Klein (2005) analyses this as an OCP effect on the feature [sonorant]. A similar phenomenon occurs for the cognate *-in-* infix in Tagalog (Schachter and Otnes 1972: 364–366), though not for *-um-*.

(14) Toba Batak /-um-/

		<i>/um/</i> ₁	<i>/um/</i> ₂	
	<i>stem</i>	<i>adjectival</i> <i>compa-</i> <i>rative</i>	<i>completive</i> <i>transitive</i>	<i>intransitive</i> ³² <i>gloss</i>
p	paddita pasak pate	up -paddita up -pasak	 m -ate	‘preacherlike’ (N72) ‘beat’ (N74) ‘die’ (N98, T117,118)
t	tibbo	t-um-ibbo		‘tall’ (N75)
b	balga bongal	ub -balga m -onjal ³³		‘big’ (N71) ‘hear’ (T117)
d	deŋgan	d-um-eŋgan		‘good’ (N64)
g	gabe gutfaŋ	g-um-abe g-um-utfaŋ		‘rich’ (T120) ‘wave violently’ (N92)
s	surut		s-um-urut	‘retreat’ (T117,118; gloss P70)
ɕ	ɕou ɕebe		ɕ-um-ou ɕ-um-ebe	‘call’ (N71) ‘move sideways’ (T118)
n	nɛaŋ	un -nɛaŋ ³⁴		‘light/lighter’ (C ₉₈ 593)
ŋ	ŋali	uŋ -ŋali		‘cold’ (N63)
r	rɔa	r-um-ɔa		‘ugly’ (N71)
h	hatop haliaŋ hutor	h-um-atop h-um-aliaŋ		‘fast’ (N82) ‘surround’ (N81)
			h-um-utur	‘be shaking’ (N100, T119)
V	uli uba	umm -uli	 um -uba ~ m -uba ³⁵	‘beautiful’ (N64) <i>no gloss</i> (T117)

tʃ, m, l *no data from these sources*

As discussed by Crowhurst (1998, 2001), a change seems to have taken place in Toba Batak since the 1950s. Percival (1981, [P] above), drawing on fieldwork from 1957 through 1977, reports that—at least for comparatives—the prefix version of /um/₁ is also possible before other consonants (e.g., *g-um-ogo* ~ *um-gogo* ‘stronger’ (p. 39)), and even preferred for certain words. However, he states that “[i]n general my informants regard the forms with the infix as ‘more correct’, and in some cases

³²Van der Tuuk states—possibly referring to both transitives and intransitives—that “[t]he infix is not used in words of which the first syllable is closed with *m*”, as in *ɕomba* (no gloss; p. 117).

³³Van der Tuuk reports that some *b*-initial stems show no reflex of the infix: *boltok* ‘become pregnant’ (p. 82, 117).

³⁴Nababan gives *un-* prefixation as the general rule for *n*-initial stems, but the example she gives is apparently exceptional in that it shows variation between prefixation and infixation: *n-um-ɛaŋ* ~ *un-nɛaŋ* ‘lighter’ (p. 62).

³⁵Van der Tuuk states that at least one verb “never loses the *u*” (*umaha* “often spelt *ummaha*”) (p. 117). Most of the examples he gives for V-initial stems have just *m-*, with no *um-/umm-* variant mentioned.

label the forms with the prefix as ‘childish’ or ‘unacceptable’ ” (p. 39). Crowhurst’s own consultations with a native speaker (C. Napitupulu) in the late 1990s indicate that—at least for comparatives and transitives—prefixation had become the norm for all consonants.

Crowhurst analyzes prefixation as driven by an alignment constraint requiring the surface form to contain a syllable consisting solely of correspondents of /um/₁ (UM \approx σ ; this is violated by infixation: *t-u.m-ib.bo*). Under her analysis of the variety shown in (14), prefixation is the default and infixation occurs only when prefixation-plus-total-assimilation would require the /m/ to assimilate in both place and nasality, as in **ug-gabe* (the innovative form of *g-um-abe*). But the prefixed data can equally be analyzed with $*_{[\text{stem} \text{ LAB } V_0 \text{ LAB}]}$, $*_{[\text{stem} \text{ NAS } V_0 \text{ NAS}]}$ \gg ALIGN, where, as above, ALIGN requires the word to begin with root material rather than affix material. (The total assimilation of *m* to a following *p* or *b* is a productive part of Toba Batak phonology—see Nababan 1981 and also Hayes 1986. The assimilation of *m* to a following *n* or η is restricted to this affix, or perhaps to affixes in general.) In order to capture the difference between the two affixes, it seems to be necessary to employ separate UNIFORMITY (anti-fusion) constraints for each affix: UNIFORMITY-/um/₁ \gg ALIGN, so that fusion is not an option for /um/₁ (*up-pasak* instead of **m-asak*). The gemination of /m/ before a vowel-initial root can be accounted for with a constraint inspired by Crowhurst’s UM \approx σ that requires the /m/ of /um/₁ to correspond to a coda consonant (though it may be split or lengthened, in violation of INTEGRITY (McCarthy and Prince 1995), so that it also corresponds to an onset consonant in the geminated case).

By contrast, /um/₂ is not subject to UNIFORMITY-/um/₁, so ALIGN and $*_{[\text{stem} \text{ LAB } V_0 \text{ LAB}]}$ can be satisfied by fusion (unfortunately there are no data for nasal-initial stems). Nor is it subject to the constraint requiring /m/ to correspond to a coda consonant, so when the root is vowel-initial it appears as a prefix *um-* or *m-*, depending on the variable ranking of ONSET and MAX-V.

In Kanakanavu (Wu 2006 [W]; Lu (in progress)), /-um-/ is metathesized and prefixed to labial- and nasal-initial stems, as in (15). If changing the infix to a prefix allows it to escape the markedness violation (although a *labial* ... *labial* or *nasal* ... *nasal* sequence is still present), Łubowicz’s (to appear) claim for Palauan that the restriction holds only of the stem domain, which includes infixes but excludes prefixes, is supported. Crowhurst’s analysis of Toba Batak would not extend to Kanakanavu, since nothing should prevent *mu-* prefixation before a non-labial, non-nasal consonant. Rather, it must be infixation that is blocked with labial- and nasal-initial stems, suggesting $*_{[\text{stem} \text{ LAB } V_0 \text{ LAB}]}$, $*_{[\text{stem} \text{ NAS } V_0 \text{ NAS}]}$ \gg ALIGN, as we proposed for Toba Batak /um/₁. In addition, ONSET must outrank LINEARITY (McCarthy and Prince’s 1995 constraint against metathesis).

(15) Kanakanavu /-um-/

<i>stem- initial C</i>	<i>stem</i>	<i>actor focus</i>	<i>gloss</i>
p	panaʔu	mu -panaʔu	‘shoot’ (W6)
t	taŋi	t-um-aŋi	‘cry’ (W7)
c	caŋu	c-um-aŋu	‘pierce’ (W7)

k	kən	k-um-ən	‘eat’ (W7)
s	sɪmən	s-um-ɪmən	‘play’ (W7)
v	vua	mu -vua	‘give’ (W6)
m	maɬəmaɬə	mu -maɬəmaɬə	‘caress by stroking’ (W6)
ɲ	ɲətɲən	mu -ɲətɲən	‘cut with a polo’ (W7)
ʔ	ʔaʔa	ʔ-um-aʔa	‘sharpen on a whetstone’ (W7)
ts, h, n, l, r		<i>no data</i>	

8 Mixed systems

Thao (Li 1976 [L]; Blust 1998a, 1998b [B_a, B_b]; Blust 2003b) has a combination of strategies when labial-initial stems are affixed with /-(u)m-/: zero realization of the affix for ϕ , and loss of stem consonant for p . Blust (1998a, p. 14) states that there is a gap for 2b , nasals, glides, and 2d : stems beginning in these consonants simply lack an actor focus form, or use a different affix or an auxiliary (and explicitly contrasts this with the zero realizations for ϕ). Blust (1998b:349), however, gives examples of zero affixation for not only for ϕ , but also for 2b , 2d , and m . The consonants n , w , and j (“y”), on the other hand, are listed as “no good example”. It is unclear whether “no good example” indicates a gap—in contrast to the ϕ , 2b , 2d , m examples—or indicates that no clear verb stems with those initial consonants were found that could, semantically, have had an actor focus form, and that thus it is impossible to tell whether there is a true gap or not. For stems beginning in l , r , δ , prefixation occurs. See Lu (2008) for an analysis of Thao infixation.

(16) Thao /-(u)m-/³⁶

<i>stem-initial C</i>	<i>stem</i>	<i>actor focus</i>	<i>gloss</i>
p	pataɸ	m -ataɸ	‘write’ (B _a 14, B _b 349)
t	tiuð	t-m-iuð	‘comb’ (B _a 13)
k	kiɸkiɸ	k-m-iɸkiɸ	‘shave’ (B _a 13)
q	qusað	q-m-usað	‘rain’ (B _a 13, B _b 349)
2b	2 bɪqsɪr	2 bɪqsɪr (<i>zero</i>)	‘seize by the throat’ (B _b 349)
		gap	(B _a 14)
2d	2 duruk	2 duruk (<i>zero</i>)	‘stab’ (B _b 349)
		gap	(B _a 14)
ϕ ($< {}^*b$ —B _a 15)	ϕ ariw	ϕ ariw (<i>zero</i>) ³⁷	‘buy’ (L233, B _a 14, B _b 349)
θ	θ anit	θ -m-anit	‘weep’ (L233, B _a 13, B _b 349)
s	sansan	s-m-ansan	‘warm by a fire’ (B _a 13)

³⁶In stems that begin with a consonant cluster, the /u/ appears, and the /m/ is assimilated to the following consonant: θ -um-piq ‘to thresh’, k -um-tun ‘to cut’. If the second consonant is a nasal, the /m/ does not surface: f -u-nuru ‘recognize’. If the first consonant is /q/ (or, in Blust 1998a: 15, /r/), the vowel surfaces as [o]: q -om-pit ‘pinch’ (Blust 1998b: 349). Li (1976) notes several cases where he has transcribed [m], but earlier fieldworkers had [um], e.g., t -m-itif (p. 233).

³⁷Li (1976: 235) lists at least one non-labial-initial “active infinitive” with zero affixation ([kaθu] ‘bring’) but there is some conflation of infinitives and imperatives in this section.

ʃ	ʃinaw	ʃ-m-inaw	‘wash, as dishes’ (L233, B _b 349)
ɬ	ɬukɬuk	ɬ-m-ukɬuk	‘hiccup’ (B _a 13)
ð	ðai	m-ðai	‘tell, advise’ (B _a 13, B _b 349)
m	mamaða	mamaða (<i>zero</i>) or m-amaða <i>gap</i>	‘threaten’ (B _b 349) (B _a 14)
l	luslus	m-luslus	‘grate or shred’ (B _a 13, B _b 349)
r	riut	m-riut	‘encircle, surround’ (B _a 13)
h	haɸuj	h-m-aɸuj	‘to chant’ (B _a 13)
n, w/β, j	<i>no good example</i>		(B _b 349)
	<i>gap</i>		(B _a 14)
V ³⁸	ara	m-ara	‘fetch’ (B _a 10, B _b 349)

The pattern described in Blust (1998b), with no gaps, can be analyzed with *_{[stemLAB V₀ LAB » UNIFORMITY} to account for *p* and for *m*, if we analyze the initial consonant of *mamaða* as a fusion of the infix /m/ and the stem /m/. Fusion with the affix is not available for other labials, *ɸ* and ²*b*: fusion of the affix and *ɸ* can be ruled out by IDENT(continuant), leaving no option for *ɸ* but deletion of the affix; and fusion of the affix and ²*b* would require either creating a glottalized ²*m*, which is not in the language’s consonant inventory, or losing the underlying glottalization. Total deletion of the affix is left as the only option, in violation of MAX. Li (1976: 232) gives a list of word-initial consonant clusters that is meant to be exhaustive; although the list is substantial (39 clusters), the voiced, non-nasal consonants ²*d*, *ð*, *l*, *r* never occur as the first member of a cluster, so general phonotactic constraints can rule out infixation after them.³⁹ For *ð*, *l*, *r*, this forces prefixation, in violation of ALIGN (which, as above, requires some stem material to be word-initial). The pre-glottalized consonant ²*d* is also restricted in its ability to appear in the second position of a cluster (*p²d* and *q²d* are the only examples),⁴⁰ suggesting that another phonotactic constraint rules out prefixation in this case, leaving no option for ²*d* but deletion of the affix.

As in the other cases of fusion, it would be possible to have an analysis in which fusion is the default but it is available only for *p* (for faithfulness reasons). It is also possible to analyze prefixation as the default here, but rule it out for phonotactic reasons with certain consonants (the only *nasal-C* onset clusters listed by Li are *mr* and *mj*).

Another mixed system seems to occur in Lolak (Danie et al. 1996, [D]),⁴¹ where /-um-/ fuses with /p/ (with contradictory data on whether the infix /u/ deletes), but

³⁸Li (1976) gives examples of [ʔ]-initial words (p. 226), but no examples of vowel-initial words, so it is possible that he would transcribe *ara* as *?ara*.

³⁹Li seems not to have collected *h-m-aɸuj*, since he states explicitly that, unlike Li et al. (1956), he did not find any instances of initial *hm*. Except for *t*, none of the consonants that can occur as the first member of a consonant cluster are found only before *m*—that is, they aren’t being counted as able to occur in the environment [_{word}__C merely by virtue of the infixed forms in (16).

⁴⁰This is harder to be certain about, though, as most consonants that can appear as the second member of an initial cluster appear in only one or two clusters.

⁴¹Thanks to a reviewer for pointing out this case.

may be prefixed to /ʔ/⁴²—we say “may be” because there is only one example given, and the prose description of /-um-/’s behavior (p. 65) does not mention any special behavior for /ʔ/, so it is possible that the example given is an error. The infix /-im-/ is prefixed to both /p/ and /d/.

(17) Lolak /-um-/

<i>stem-initial C</i>	<i>stem</i>	<i>/-um-/ future</i>	<i>/-im-/ past</i>	<i>gloss</i>
p	puti?	um -uti? (D66) m -uti? (D131)		‘whiten’
	połok		im -połok (D67)	‘shorten’
t	tindok	t-um-indok (D65)	t-im-indok (D132)	‘stand’
k	kosij	k-um-osij (D65)	k-im-osij (D66)	‘laugh’
d	deder	d-um-eder (D66)		‘recline’
	dałom		in -dałom (D67)	‘deepen’
g	gotup	g-um-otup (D65)	g-im-otup (D67)	‘burst’
s	suot	s-um-uot (D66)	s-im-uot (D67)	‘enter’
ʔ	ʔanduk	mu-ʔanduk (D66)		‘jump’
	ʔitu?		ʔ-im-itu? (D66)	‘sit down’
r	rogi	r-um-ogi (D131)	<i>no data</i>	‘be angry’
ʔ ⁴³	ʔobom	ʔ-um-obom (D66)		‘lie’
	ʔintok		ʔ-im-intok (D67)	‘diminish’
b, m, n, ŋ, l, w, j, h	<i>no data</i>			

9 Conclusion

The sections above have presented a variety of “repairs” for potential *labial...labial* sequences created by *um*-infixation in Austronesian languages. We have seen that avoidance of these sequences seems to apply only within the root-and-infix domain,⁴⁴ and only in derived environments. The table in (18) shows that many of the languages can be characterized as avoiding the structure [_{stem}LAB V₀ LAB, a stem-initial labial with the next consonant being labial, and that the array of repairs is diverse, both among languages where the infix is *Vm* (so that the two labials would be separated by a vowel) and those with *m* (so that the two labials would be strictly adjacent). We also see one case of a more limited constraint, against a stem-initial *sonorant* labial when the next consonant is labial (Tagalog), and cases of more general constraints: Palauan, where a second labial consonant anywhere in the stem is forbidden; Tjuabar Paiwan where the restriction holds even if the first labial consonant is not stem-initial—in these two languages, the more-specific constraint [_{stem}LAB V₀ LAB may also be

⁴²A lateral flap (Danie et al. 1996: 22).

⁴³Transcriptions in Danie et al.’s “Data Fonologi” appendix, pp. 313–319, indicate that words broadly transcribed with an initial vowel are /ʔ/-initial.

⁴⁴For a different view, see the Klein (2005) analysis of Tagalog, Chamorro, and Toba Batak—he proposes that the markedness constraints in question apply only to the first bimoraic foot of the stem. However, as Klein notes, the footing required conflicts with the footing needed to account for stress.

ranked high—and Kavalan, where the restriction holds even across the prefix-stem boundary (as it does in Palauan for a different affix).

(18)	<i>Phonotactic</i>	<i>Reaction</i>	<i>/Vm/ languages</i>	<i>/m/ languages</i>
	*[_{stem} LAB V ₀ LAB	fusion	T. Murut, M. Atayal, Muna, Toba Batak /um/ ₂ , S. Manobo, A. Tagbanwa, Kimaragang, Sediq	S. Blaan, Thao
		zero affixation gap prefixation	Muna N. Acehnese, Toba Batak /um/ ₁ , Kanakanavu	Thao Thao
		infix dissim. stem dissim.	L. Kalinga, N. Acehnese	K. Paiwan
		tolerance	Tagalog	
	*[_{stem} SONLAB V ₀ LAB	gap	Tagalog	
	*[_{stem} LAB... LAB	fusion		Palauan
		infix dissim.		Palauan
	*LAB(ə)LAB	infix dissim.	Tjuabar Paiwan	
	*LAB VOICEDLAB	loss of infix nasality		Kavalan
	*[_{stem} NAS V ₀ NAS	zero affixation prefixation	Muna Toba Batak /um/ ₁ , Kanakanavu	
	*[_{stem} SON V ₀ SON	prefixation	N. Acehnese	S. Blaan
	non-OCP	fusion	Kapampangan	

We saw that it was possible to analyze fusion as not driven by markedness constraints but instead as the default, allowed by faithfulness to apply only to labial consonants. We also saw that it was possible to treat prefixation as the default in some cases: Toba Batak (where, as Crowhurst 1998, 2001 shows, infixation is being eliminated in favor of prefixation) and Thao. As Crowhurst suggests, prefixation may be especially vulnerable to reanalysis if infixation is not a primitive type of morphology but rather an unusual strategy for obeying markedness constraints (e.g., NOCODA in Prince and Smolensky 1993/2004). Prefixation to avoid OCP-labial sets the stage for prefixation to avoid other OCP constraints, encouraging a reanalysis that leads to further restriction of infixation. On the other hand, we saw that gaps, zero affixation, some prefixation (Kanakanavu, Northern Acehnese), infix dissimilation, and stem dissimilation required an OCP-driven analysis.

The genetic affiliations of the languages discussed here are shown in (19), using an amalgam of the groupings in Grimes et al. (1995) and Tryon (1995a, 1995b) (as Tryon discusses, they are not free of controversy). It seems that *um*'s behavior bears little relation to genetic affiliation, suggesting independent innovation of the various repairs.

(19)

			Language	Processes
Austronesian	Malayo-Polynesian	Atayalic	Mayrinax Atayal	fusion
			Sediq	fusion
		Rukai-Tsouic	Kanakanavu	prefixation
		Paiwanic	Paiwan	infix dissimilation, stem-internal effects
			Thao	fusion, zero, gap
			Kavalan	loss of nasality
		N. Philippine	Limos Kalinga	stem dissimilation
			Kapampangan	fusion
		Meso-Philippine	Tagalog	gap
			Aborlan Tagbanwa	fusion
		S. Philippine	Sarangani Manobo	fusion, zero
			Lolak	fusion, prefixation
	W. Malayo-Polynesian	S. Mindanao	Sarangani Blaauw	fusion
		Borneo	Timugon Murut	fusion
			Kimaragang	fusion
		Sundic	N. Acehnese	stem dissimilation, prefixation
			Toba Batak	fusion, prefixation
		Sulawesi	Muna	fusion, zero
		Palauan	Palauan	fusion, infix dissim., stem-internal effects

How well does this diversity of repairs fit with current responses to the too-many-solutions problem? Steriade (2001/2008) proposes that faithfulness constraints are, by default, ranked so that changes creating a perceptually greater departure are penalized more heavily. That is, changes should be perceptually minimal (e.g., final devoicing rather than epenthesis), unless of course the perceptually minimal change is ruled out by some markedness constraint. So why do “solutions” to *labial*... *labial* sequences differ from language to language? In some cases, ranking of non-faithfulness constraints could explain cross-linguistic differences. For example, a high ranking of a constraint requiring the *um* morpheme to be realized could rule out a zero realization in most languages, as could a constraint against paradigm gaps (e.g., Prince and Smolensky’s 1993/2004 MPARSE, elaborated in Wolf and McCarthy 2005 (to appear); Rice’s 2005 MAX{CAT} constraints), or a lenient CONTROL component under Orgun and Sprouse’s 1999 theory. But the difference between infix dissimilation and stem dissimilation, for example, shouldn’t be attributable to markedness differences—in both cases, the result is a phonotactically unproblematic string. This is true whether the change is the relatively subtle one from *m* to *n*, as in Tjuabar Paiwan /p-əm-ajsu/ → [p-ən-ajsu], or the presumably more salient change from initial *p* to *k*, as in Limos Kalinga /p-um-ija/ → [k-um-ija]. If only faithfulness is at stake, we expect the Paiwan pattern rather than the Limos Kalinga. Perhaps the explanation is that pressures against potential neutralization in stems and affixes exist, and are ranked differently in different languages. Another possibility, however, is that *all*

possible repairs for *labial* ... *labial* sequences involve substantial departures from the faithful form (compared to final devoicing and the like), and that in such cases where no perceptually close repair is available, the choice of repair is arbitrary and language-specific.⁴⁵

This explanation of diversity (that no minor repair is possible) does not extend, however, to the two cases mentioned in the introduction. The obvious solution to Pater's (1999) *NC constraint against a nasal followed by a voiceless obstruent is to voice the obstruent: *mp* → *mb*. Hayes and Stivers (2000) even found partial postnasal voicing in English, a language that generally tolerates *NC violations—and yet Pater shows that postnasal voicing is just one of many strategies used. Similarly, the obvious solution to Kennedy's (2003) constraint against initial geminates is degemination, but that is only one of several strategies that he documents.

An understanding of what causes repairs to be diverse or uniform will require determining what the diverse-repairs (or uniform-repairs) cases have in common. Blumenfeld (2006) places restrictions on how prosodic and segmental constraints and repairs can interact, for example. Within the realm of purely segmental phenomena, we might consider such factors as:⁴⁶

- (i) The origin of the markedness constraint: Is it driven by articulatory considerations, as Hayes and Stivers suggest for *NC, by perceptual difficulties, or, as some have suggested for OCP constraints, by motor planning difficulties (Frisch 2004)?
- (ii) The formal complexity of the markedness constraint: How long a string must be inspected to determine if there is a violation? Is it sensitive to morphological information or other hidden structure? How many features are involved?
- (iii) The nature of the changes available: Is one change perceptually closer to the original than the others (and if so, does it achieve the status of “only solution” by falling below some threshold of perceptual distance, or must the difference between the closest change and the next-closest fall above some threshold)? Does one change affect fewer segments, fewer features, or less-important features? If each change is formulated as a rule, does one change have a simpler structural description?

While a squib about one affix and one type of markedness cannot determine the crucial factor(s), we note that all the grammars here must make reference to morphological information or other hidden structure (e.g., the left boundary of the stem). Although *NC and *INITIALGEMINATE in principle do not require such information, the cases that Pater and Kennedy examine do involve particular morphological boundaries. By contrast, avoidance of a final voiced obstruent (the Steriade/Lombardi case) is generally across the board.

Acknowledgements We thank Bruce Hayes, Russ Schuh, Donca Steriade, and the reviewers and associate editor for all their comments and ideas.

⁴⁵Thanks to Bruce Hayes for generating this idea; he should not be held responsible for its viability.

⁴⁶We thank a reviewer for suggesting these possibilities.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Abrams, Norman. 1961. Word base classes in Bilaan. *Lingua* 10: 391–402.
- Abrams, Norman. 1970. *Bilaan morphology. Papers in Philippine linguistics, Vol. 3. Vol. 24 of Pacific Linguistics Series A*, 1–62. Canberra: Australian National University.
- Baković, Eric. 2007. Local assimilation and constraint interaction. In *The Cambridge handbook of phonology*, ed. Paul De Lacy, 335–352. Cambridge: Cambridge University Press.
- Balilo, Zsa-Zsa. 2008. Kapampangan um- verbs. Ms., UCLA.
- Blevins, Juliette. 2004. *Evolutionary phonology: The emergence of sound patterns*. Cambridge: Cambridge University Press.
- Blumenfeld, Lev. 2006. Constraints on phonological interactions. Ph.D. dissertation, Stanford University.
- Blust, Robert. 1998a. Some problems in Thao phonology. In *Selected papers from the second international symposium on language in Taiwan (ISOLIT-II)*, ed. Shuanfan Huang, 1–20. Taipei: Crane.
- Blust, Robert. 1998b. A note on the Thao patient focus perfective. *Oceanic Linguistics* 37: 346–353.
- Blust, Robert. 2003a. Three notes on early Austronesian morphology. *Oceanic Linguistics* 42: 438–478.
- Blust, Robert. 2003b. *Thao dictionary*. Taipei: Institute of Linguistics Preparatory Office, Academia Sinica.
- Blust, Robert. 2004. Austronesian nasal substitution: a survey. *Oceanic Linguistics* 43: 73–148.
- Carter, Allyson. 2000. Featural morphology: evidence from Muna irrealis affixation. In *Coyote papers: Working papers in linguistics from A-Z 10*, eds. Amy Fountain, Sean Hendricks, Sachiko Ohno, Mizuki Miyashita, and Debbie Cole, 1–17. Tucson: University of Arizona Linguistics Circle.
- Crowhurst, Megan. 1998. Um infixation and prefixation in Toba Batak. *Language* 74: 590–604.
- Crowhurst, Megan. 2001. Coda conditions and um infixation in Toba Batak. *Lingua* 111: 561–590.
- Danie, J. Akun, A.B.G. Rattu, J. Ranguban, H.N. Meruntu, and N. Hiarey. 1996. *Struktur Bahasa Lolak*. Jakarta: Pusa Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan.
- De Lacy, Paul. 2003. Fixed ranking and the ‘Too many solutions’ problem. CASTL Kick-off conference handout.
- del Corro, Anicia. 1980. *Kapampangan morphophonemics. The Archive: a journal devoted to the study of Philippine languages and dialects*. Diliman, Quezon City: Cecilio Lopez Archives of Philippine Languages and the Philippine Linguistics Circle, University of the Philippines.
- DuBois, Carl D. 1976. *Sarangani Manobo: an introductory guide*. Manila: Linguistic Society of the Philippines.
- Durie, Mark. 1985. *A grammar of Acehnese on the basis of a dialect of North Aceh*. Dordrecht: Foris.
- English, Leo. 1986. *Tagalog-English dictionary*. Manila: Congregation of the Most Holy Redeemer. Distributed by National Book Store.
- Ferreirinho, Naomi. 1993. *Selected topics in the grammar of Limos Kalinga, the Philippines*. Vol. 109 of *Pacific Linguistics Series B*. Canberra: Australian National University.
- Ferrell, Raleigh. 1982. *Paiwan dictionary*. Vol. 73 of *Pacific Linguistics Series C*. Canberra: Australian National University.
- Finer, Daniel L. 1986. Reduplication and verbal morphology in Palauan. *The Linguistic Review* 6: 99–130.
- Flora, Jo-Ann. 1974. Palauan phonology and morphology. Ph.D. Dissertation, University of California, San Diego.
- Forman, Michael L. 1971. *Kapampangan grammar notes*. Honolulu: University of Hawaii Press.
- Frisch, Stefan. 1996. Similarity and frequency in phonology. Ph.D. Dissertation, Northwestern University.
- Frisch, Stefan. 2004. Language processing and segmental OCP effects. In *Phonetically-based phonology*, eds. Bruce Hayes, Robert Kirchner, and Donca Steriade, 346–371. Cambridge: Cambridge University Press.
- Frisch, Stefan, Janet Pierrehumbert, and Michael Broe. 2004. Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22: 179–228.
- Green, Peter. 1979. Co-existent aspect-marking phenomena in Aborlan Tagbanwa. In *Papers in Philippine linguistics, Vol. 9*, eds. E. Clay Johnston, Hartmut Wiens, Jo Ann Gault, Peter Green, Bruce Grayden, and Jeanette Witucki, Vol. 50 of *Pacific Linguistics Series A*, 69–90. Canberra: Australian National University.

- Grimes, Barbara, Joseph Grimes, Malcolm Ross, Charles Grimes, and Darrell Tryon. 1995. Listing of Austronesian languages. In *Comparative Austronesian dictionary: An introduction to Austronesian studies, Part 1: Fascicle 1*, ed. Darrell T. Tryon, 121–279. Berlin/New York: de Gruyter.
- Hayes, Bruce. 1986. Assimilation as spreading in Toba Batak. *Linguistic Inquiry* 17: 467–499.
- Hayes, Bruce, and Tanya Stivers. 2000. Postnasal voicing. Ms., UCLA.
- Ho, Dah-an. 1977. The phonological system of Butanglu: A Paiwan dialect. *Bulletin of the Institute of History and Philology Academia Sinica* 48(4): 595–618.
- Ho, Dah-an. 1978. A preliminary comparative study of five Paiwan dialects. *Bulletin of the Institute of History and Philology Academia Sinica* 49(4): 565–681.
- Ho, Dah-an. 1995. Paiwan. In *Comparative Austronesian dictionary: an introduction to Austronesian studies. Part 1: Fascicle 1*, ed. Darrell T. Tryon, 307–314. Berlin/New York: de Gruyter.
- Huang, Lillian. 2000. Verb classification in Mayrinax Atayal. *Oceanic Linguistics* 39: 364–390.
- Itô, Junko, and Armin Mester. 1995a. Japanese phonology. In *The handbook of phonological theory*, ed. John Goldsmith, 817–838. Cambridge/Oxford: Blackwell.
- Itô, Junko, and Armin Mester. 1995b. The core-periphery structure of the lexicon and constraints on reranking. In *University of Massachusetts occasional papers in linguistics [UMOP]*, eds. Jill Beckman, Suzanne Urbanczyk, and Laura Walsh, Vol. 18 of *Papers in Optimality Theory*, 181–209. Amherst: GLSA.
- Itô, Junko, and Armin Mester. 1999. The structure of the phonological lexicon. In *The handbook of Japanese linguistics*, ed. Natsuko Tsujimura, 62–100. Malden/Oxford: Blackwell.
- Josephs, Lewis. 1975. *Palauan reference grammar*. Honolulu: University of Hawaii Press.
- Josephs, Lewis. 1990. *New Palauan-English dictionary*. Honolulu: University of Hawaii Press.
- Kennedy, Robert. 2003. *Confluence in phonology: evidence from Micronesian reduplication*. Ph.D. dissertation, University of Arizona.
- Kisseberth, Charles. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1: 291–306.
- Klein, Thomas. 2005. Infixation and segmental constraint effects: UM and IN in Tagalog, Chamorro, and Toba Batak. *Lingua* 115: 959–995.
- Kroeger, Paul. 1988. Verbal focus in Kimaragang. In *Papers in Western Austronesian linguistics, Vol. 3*, ed. Hein Steinhauer, Vol. 78 of *Pacific Linguistics Series A*, 217–240. Canberra: Australian National University.
- Li, Fang Keui, Chi-lu Chen, and Mei-chün Tang. 1956. Notes on the Thao language. *Bulletin of the Department of Archeology and Anthropology, National Taiwan University* 7: 23–51.
- Li, Paul Jen-kuai. 1976. Thao phonology. *Bulletin of the Institute of History and Philology Academia Sinica* 47(2): 219–244.
- Li, Paul Jen-kuai. 1991. Vowel deletion and vowel assimilation in Sediq. In *Currents in Pacific linguistics: Papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, ed. Robert Blust, Vol. 117 of *Pacific Linguistics Series C*, 163–169. Canberra: Australian National University.
- Li, Paul Jen-kuai. 1995. Atayal. In *Comparative Austronesian dictionary: an introduction to Austronesian studies. Part 1: Fascicle 1*, ed. Darrell T. Tryon, 283–292. Berlin/New York: de Gruyter.
- Li, Paul Jen-kuai, and Shigeru Tsuchida. 2006. *Kavalan dictionary*. Taipei: Institute of Linguistics, Academia Sinica.
- Lombardi, Linda. 1995/2001. Why place and voice are different: constraint-specific alternations in optimality theory. In *Segmental phonology in optimality theory: Constraints and representations*, ed. Linda Lombardi, 13–25. Cambridge: Cambridge University Press. Originally circulated as ms., University of Maryland.
- Lu, Yu-an. 2005. *Mayrinax phonology: with special reference to UM affixation*. MA thesis, National Tsing Hua University.
- Lu, Yu-an. 2008. UM infixation in Thao. Ms., SUNY at Stony Brook, New York.
- Lu, Yu-an. In progress. UM infixation and paradigmatic gaps in Thao. Ms., SUNY at Stony Brook, New York.
- Łubowicz, Anna. To appear. Infixation as morpheme absorption. In *Phonological argumentation: Essays on evidence and motivation*, ed. Steven Parker. London: Equinox.
- McCarthy, John. 1979. Formal problems in Semitic phonology and morphology. Ph.D. dissertation, MIT.
- McCarthy, John. 2002. *A thematic guide to optimality theory*. Cambridge: Cambridge University Press.
- McCarthy, John, and Alan Prince. 1993. Generalized alignment. In *Yearbook of morphology*, eds. Geert Booij and Jaap van Marle, 79–153. Dordrecht: Kluwer.
- McCarthy, John, and Alan Prince. 1995. Faithfulness and reduplicative identity. In *University of Massachusetts occasional papers in linguistics 18: Papers in optimality theory*, eds. Jill Beckman, Suzanne Urbanczyk, and Laura Walsh Dickey, 249–384. Amherst: GLSA.

- McLachlin, Betty, and Barbara Blackburn. 1968. Verbal clauses of Sarangani Bilaan. *Asian Studies* 6: 108–128.
- Mester, R. Armin. 1986. Studies in tier structure. Ph.D. dissertation, University of Massachusetts, Amherst.
- Nababan, P.W.J. 1981. *A grammar of Toba-Batak. Materials in languages of Indonesia No. 6*. Vol. 37 of *Pacific Linguistics Series D*. Canberra: Australian National University.
- Orgun, Orhan, and Ronald Sprouse. 1999. From MPARSE to CONTROL: deriving ungrammaticality. *Phonology* 16: 191–224.
- Pater, Joe. 1999. Austronesian Nasal substitution and other NC effects. In *The Prosody morphology interface*, eds. René Kager, Harry van der Hulst, and Wim Zonneveld, 310–343. Cambridge: Cambridge University Press.
- Pater, Joe. 2001. Austronesian nasal substitution revisited: what's wrong with *NÇ (and what's not). In *Segmental phonology in optimality theory: Constraints and representations*, ed. Linda Lombardi, 159–182. Cambridge: Cambridge University Press.
- Pater, Joe. 2003. Balantak metathesis and theories of possible repair. Ms., University of Massachusetts, Amherst.
- Percival, W.K. 1981. *A grammar of the urbanised Toba-Batak of Medan*. Vol. 76 of *Pacific Linguistics Series B*. Canberra: Australian National University.
- Prentice, D.J. 1971. *The Murut languages of Sabah*. Vol. 18 of *Pacific Linguistics Series C*. Canberra: Australian National University.
- Prince, Alan, and Paul Smolensky. 1993/2004. *Optimality theory: Constraint interaction in generative grammar*. Malden: Blackwell. Originally circulated as Technical report CU-CS-696-93, Department of Computer Science, University of Colorado at Boulder, and Technical report TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ.
- Rhea, Mary. 1972. Prefocus and verbal orientation in Sarangani Bilaan. *Philippine Journal of Linguistics* 3: 35–42.
- Rhea, Mary. 1995. Sarangani Blaauw. In *Comparative Austronesian dictionary: an introduction to Austronesian studies. Part 1: Fascicle 1*, ed. Darrell T. Tryon, 369–373. Berlin/New York: de Gruyter.
- Rice, Curt. 2005. Optimal gaps in optimal paradigms. *Catalan Journal of Linguistics* 4: 155–170.
- Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice-marking. In *The history and typology of western Austronesian voice systems*, eds. Fay Wouk and Malcolm Ross, Vol. 518 of *Pacific Linguistics*, 17–62. Canberra: Australian National University.
- Schachter, Paul, and Fe Otanes. 1972. *Tagalog reference grammar*. Berkeley: University of California Press.
- Steriade, Donca. 2001/2008. The phonology of perceptibility effects: the P-Map and its consequences for constraint organization. In *The nature of the word: studies in honor of Paul Kiparsky*, eds. Sharon Inkelas and Kristin Hanson, 151–180. Cambridge: MIT Press. Originally circulated as ms. UCLA/MIT.
- Tryon, Darrell. 1995a. The Austronesian languages. In *Comparative Austronesian dictionary: an introduction to Austronesian studies. Part 1: Fascicle 1*, ed. Darrell T. Tryon, 5–44. Berlin/New York: de Gruyter.
- Tryon, Darrell. 1995b. Proto-Austronesian and the major Austronesian subgroups. In *The Austronesians: Historical and comparative perspectives*, eds. Peter Bellwood, James J. Fox, Darrell Tryon, 17–38. Canberra: Anthropology, Research School of Pacific and Asian Studies.
- Tseng, Meylysa. 2003. Reduplication as affixation in Paiwan. MA thesis, National Chung Cheng University, Taiwan.
- van den Berg, René. 1989. *A grammar of the Muna language*. Dordrecht: Foris.
- van den Berg, René. 1991. Muna historical phonology. In *Studies in Sulawesi linguistics, part II*, ed. James N. Sneddon, Vol. 33 of *NUSA: Linguistic Studies of Indonesian and Other Languages in Indonesia*, 3–28. Jakarta: Universitas Katolik Indonesia Atma Jaya.
- van der Tuuk, H.N. 1867/1971. *A grammar of Toba Batak*. The Hague: Martinus Nijhoff. Translation, with incorporation of van der Tuuk's manuscript notes and corrections, of *Tobasche spraakunst*.
- van Oostendorp, Marc, and Ben Hermans. 2007. Solving the too-many-solutions problem. Handout of talk given at the Meertens Institute, retrieved from www.vanoostendorp.nl.
- Wiens, Hartmut. 1979. The semantic function of focus affixes in Limos Kalinga. In *Papers in Philippine linguistics No. 9*, eds. E. Clay Johnston, Hartmut Wiens, Jo Ann Gault, Peter Green, Bruce Grayden, and Jeanette Witucki, Vol. 50 of *Pacific Linguistics Series A*, 19–47. Canberra: Australian National University.

- Wilson, Colin. 2001. Consonant cluster neutralisation and targeted constraints. *Phonology* 18: 147–197.
- Wolf, Matthew, and John McCarthy. 2005/to appear. Less than zero: Correspondence and the null output. In *Modeling ungrammaticality in optimality theory*, ed. Curt Rice. London: Equinox. Previously circulated as ms., ROA-722.
- Wu, Chun-Ming. 2006. UM affixation in Kanakanavu. Ms., National Tsing Hua University, Taiwan.
- Yang, Hsiu-fang. 1976. The phonological structure of the Paran dialect of Sediq. *Bulletin of the Institute of History and Philology Academia Sinica* 47(4): 611–706.
- Yip, Moira. 1989. Feature geometry and co-occurrence restrictions. *Phonology* 6: 349–374.
- Yip, Moira. 2007. Tone. In *The Cambridge handbook of phonological theory*, ed. Paul De Lacy, 229–251. Cambridge: Cambridge University Press.
- Zuraw, Kie. 2002. Aggressive reduplication. *Phonology* 19: 395–439.